International Journal of Computer Science Engineering
and Information Technology Research (IJCSEITR)
ISSN (P): 2249–6831; ISSN (E): 2249–7943
Vol. 12, Issue 1, Jun 2022, 115–122

SUPPLY CHAIN DEMAND FORECASTING USING APPLIED MACHINE LEARNING AND FEATURE ENGINEERING

SANDEEP JALA[1], NEETI ADHIA[2], MANUSHI KOTHARI[3], DRUMIL JOSHI[4] & RANJUSHREE PAL[5]

[1,2,4,5]Electronics and Telecommunication Engineering, Dwarkadas J. Sanghvi College of Engineering, Maharashtra, India

[3]Computer Engineering, Thakur College of Engineering and Technology, Maharashtra, India

ABSTRACT

Accurate forecasting of the demand of fast-moving consumer goods is a competitive factor for manufacturers and retailers, especially in the fields of fashion, technology, and foods. This experimental research highlights the benefits of Machine Learning in predicting the sale of shorter shelf life and more volatile products, as it exceeds the level of accuracy of traditional mathematical strategies and, consequently it improves innovation value across the supply chain in the marketplace with main motive to improve consumer access and gross profit. This paper reviews existing machine learning methods for predicting demand of food sales. It discusses important design decisions of a data scientist working on food marketing predictions, such as the volatile sales data, the different inputs used to predict sales that represent the diversity of product sales. In addition, it updates the machine learning algorithms used in food sales forecasts and the appropriate steps to check your accuracy. Finally, it discusses the major challenges and learning opportunities of the machine used in the food marketing industry.

KEYWORDS: Forecasting, Supply chain, retailers, Machine learning, Consumer good.

Received: Apr 05, 2022; **Accepted**: Apr 25, 2022; **Published**: May 10, 2022; **Paper Id**: IJCSEITRJUN202214

Original Article

INTRODUCTION

Good cuisine is thought to be the cornerstone of genuine happiness. Providing high-quality food necessitates a steady supply of products for consumers to buy. Predicting a product's demand is a significant and necessary occurrence for a salesperson in terms of time and money. When it comes to food products, there are numerous aspects to consider, like price, popularity, flavour, occupancy (space), and so on. With an expanding number of elements and consumers, forecasting demand gets more complex. Predicting the number of products to be purchased and prepared is a crucial task in the pursuit of sustainable development. It's quite impossible to forecast the quantity of orders that will be placed in a given day.

A bad prediction could result in purchasing and cooking less food, resulting in a scarcity, or purchasing and preparing more food, resulting in food waste. As a result of the uncertainty and swings in consumer demand and tastes, weather changes, and price changes, anticipating exact demand is difficult. All of this is affected by seasonal fluctuations, as some meals are only available for a limited time. As a result, seasonal swings in orders make it difficult to forecast dips and surges in orders. We are exploring how to estimate the demand for food in the future in order to solve such relevant challenges.

We investigated at food forecasting methods to determine, among other things, food demand in a city, popular cuisine trending and sold in a city, maximum sales in a branch, product cost, and number of orders. This study describes a system for forecasting approaches that uses Machine Learning algorithms and statistical analysis.

LITERATURE SURVEY

Existing machine learning algorithms for food sales prediction were utilized to explore key design decisions made by data analysts working in the field of food sales, such as the temporal granularity of sales data and the input variables to employ for predicting sales output variables. It looks at how to assess the accuracy of food sales forecasting.

POS data is used to provide demand predictions at restaurants using Machine Learning and mathematical data. Various components, such as store locations, weather, and events, should be considered when making actual business demand estimates. As a result, using machine learning, create a model that predicts the need to mix and match the above-mentioned data. As part of the study, demand forecasting methods were used to increase accuracy by combining internal data such as POS data with external data from ubiquitous data such as weather, events, and so on. As a means of predicting demand, Bayesian Linear Regression, Boosted Decision Tree Regression, Decision Forest Regression, and Stepwise method were used. The prediction levels of the Bayesian, Decision, and Stepwise methods were similar, while the Boosted prediction rate was slightly lower. Estimates of any store exceeded about 85 percent.

One of the most important features of a supply chain is demand forecasting. Its goal was to improve stock rates, reduce costs, and improve sales, profitability, and customer loyalty. A variety of comparisons and comprehensive assessments show that the system for predicting a need exceeds current studies. Unlike previous research, the proposed predictive system includes vector retrieval, in-depth reading model, and a novel compilation method to ensure maximum accuracy.

The demand forecasting model is constructed using nine different time series methods, a support vet reversal algorithm, and a DL method. We aim to increase the breadth of features in the future by accessing data from other sources such as economic research, shopping trends, social media, social events, and demographic based data. In-depth learning contributions from a variety of new data sources can be seen. Further investigation of hyper parameters of the in-depth learning algorithm can be done.

We also aim to use other in-depth learning strategies such as learning algorithms, such as convolutional neural networks, duplicate neural networks, and deep neural networks. In addition, we want to use heuristic methods such as MBO (Migrating Birds Optimization) and other comparable algorithms to improve other coefficients / weights that have been strongly established by trial and error, such as taking 30% of the most effective methods in our existing system.

Considering our research topic, it attempts to utilise Machine Learning Algorithm to estimate upcoming food production.

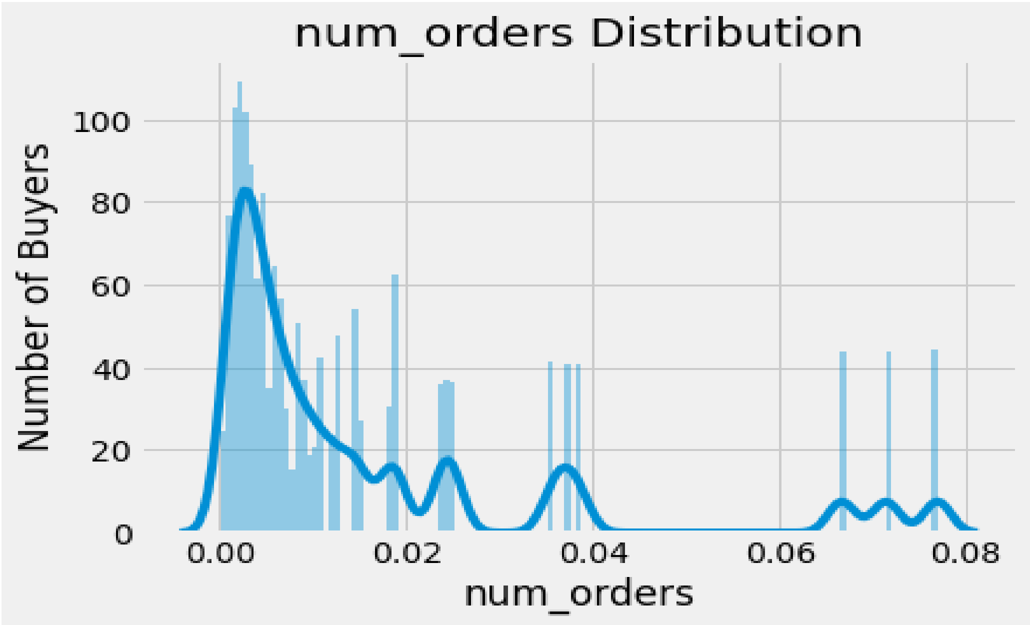

Figure 1: Distribution Plot with Number of Orders and Number of Buyers

METHODOLOGY

Exploratory Data Analysis

Analyse exploratory data EDA (Exploratory Data Analysis) is a method of analysing a dataset to find anomalies, patterns, and trends. EDA aids in gaining a general understanding and speculation of the dataset. It entails requesting summary statistics for numerical data and constructing graphs to aid in the visualisation and analysis of the data. Panda's profiling is an open-source Python package that we used. Pandas profiling makes use of the Data Frame 'df.profile report()' in Pandas to perform a quick data analysis.

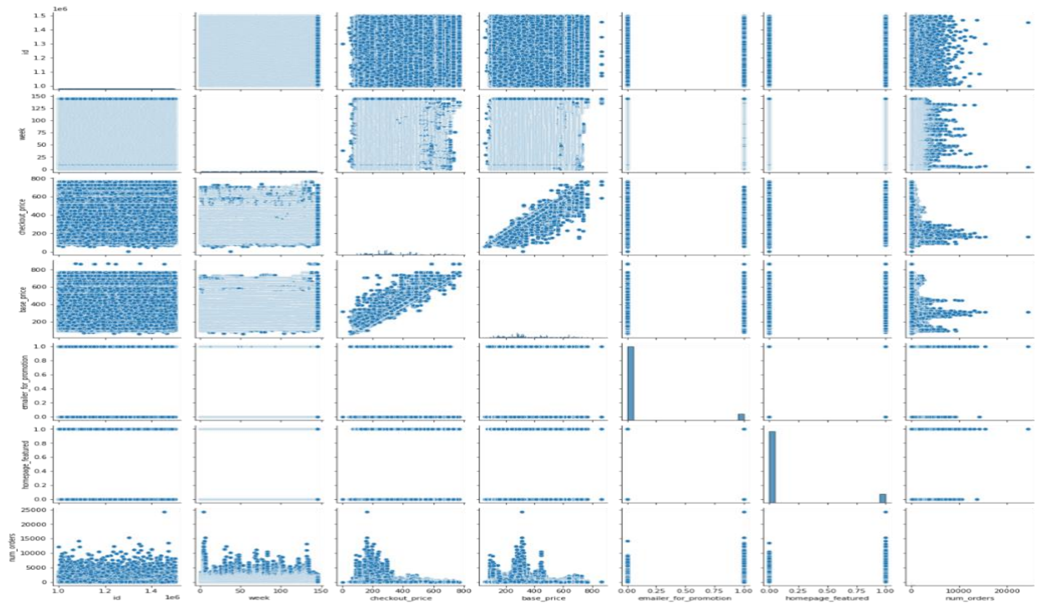

Figure 2: Pair Plot Showing Relation Between Every Variable

A heatmap is a graphical representation in which each character value of a matrix is represented by a colour for easier pattern recognition. The correlation between the several independent variables is visualised using a heat map.

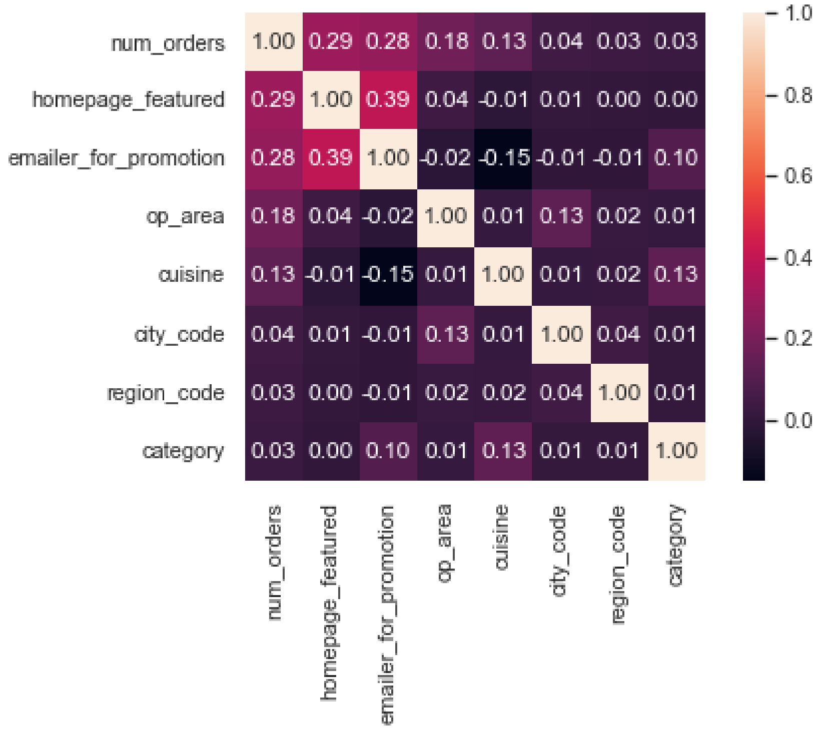

Figure 3: Heatmap with Different Features

Applied Machine Learning Algorithms

Lasso

The term "LASSO" is an acronym for Least Absolute Shrinkage and Selection Device. It is a formula for familiarizing data models and selecting features. Lasso regression is a type of linear regression that uses shrinkage. The term "shrinkage" refers to the process of reducing the amount of data to a single number, as a description. Simple, small models are popular in the lasso method (i.e. models with fewer limits). This type of retrospective is suitable for models with multiple multicollinearities or if you wish to automate the model selection process, such as dynamic selection and parameter removal.

The L1 rating is used in Lasso retrieval, and adds a fine equal to the total value of the coils. This type of adaptation can lead to smaller models with fewer coefficient; some coefficient may be zero, and the model may be excluded. Larger penalties offer coefficient values close to zero, which is great for making simple models.

$$\text{Cost(B)}=\left(\frac{1}{2n}\right)\sum_{i=1}^{n}\left(y_i-\sum_j x_{ij}B_J\right)^2+\lambda\sum_{J=1}^{p}\left|B_J\right|$$

Where λ is the amount of shrinkage.

Elastic Net

The Elastic net is a standard linear regression that combines two fine functions, L1 and L2 fine functions. The strategy incorporates lasso retrieval methods and retraction methods by learning from their mistakes to improve the mathematical model.

The elastic net approach overcomes lasso barriers, as when high-altitude data requires only a few samples. The elastic net approach allows the "n" variable to be inserted until the space is filled. If the variables are closely related groups, the lasso will usually select one from each group and ignore the others.

The stretch net includes a quadratic expression (|| β || 2) in the penalty for overcoming lasso limits, which when used alone becomes a ridge regression.

$$\text{Cost(B)}= \left(\frac{1}{2n}\right)\sum_{i=1}^{n}(y_i - x_i B)^2 + \lambda\left(\frac{1-\alpha}{2}\sum_{J=1}^{P} B_j{}^2 + \alpha\sum_{J=1}^{p}\left|B_J\right|\right)$$

where α is the mixing parameter between ridge (α = 0) and lasso (α = 1).

XGBoost

XGBoost stands for extreme Gradient Boosting. XGBoost is a faster algorithm compared to other algorithms due to its compact and distributed computer. Gradient boosting is a type of machine learning method that can be used to solve the challenges of segmentation or deceleration. XGBoost was created with your careful consideration of both system configuration and machine learning methods. The purpose of this library is to push computers to their limits in terms of computer terms in order to build a measurable, portable, and accurate library.

$$\text{L}(ø) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega f_k$$

Decision Tree

Decision Tree is a simple supervised learning algorithm. Used for troubleshooting planning and retrieval. Decisions to be made or tests to be performed by making a training model used to predict the class or number of targeted variables by reading simple decision rules taken from previous data (training data). Therefore, the decision tree is a metaphor for finding possible solutions to different problems depending on the circumstances. The decision tree has a root node that is also divided into different nodes: the area of the decision, and the leaf node. Decision Nodes are used for decision making and have many branches and Leaf nodes represent the effect of those decisions and have no other branches. Decisions or tests are made on the basis of the data provided. The Gini Indicator is used to determine whether binary separation is required in the database. The total value of the Gini index is 0 and the worst is 0.5 (in two-phase problems). The Gini Index is calculated using the following arithmetic:

$$\text{Gini(D)} = 1 - \sum_{i=1}^{m} P_i^2$$

where Pi denotes the likelihood that a tuple in D belongs to Ci.

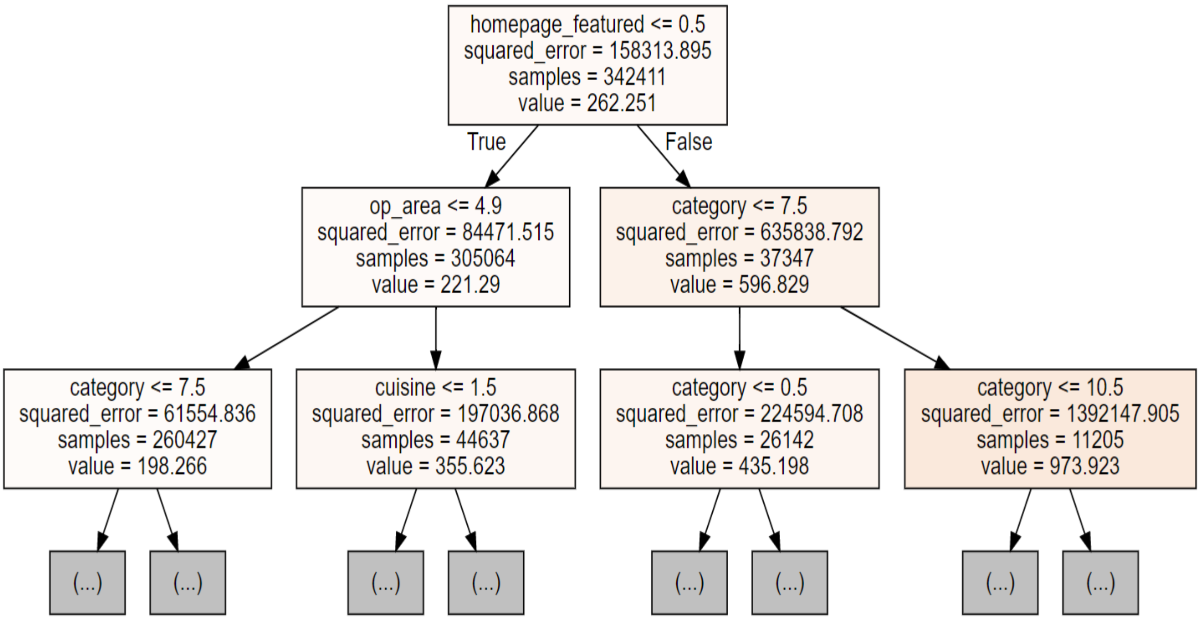

Figure 5: Decision Tree with Depth=2

Random Forest

Random Forest helps logically segregate various possible answers or outcomes in respect of a given set of information and analyses all such possible outcomes to enhance predictive accuracy of such given information. It is based on collective learning and grouping various possible decisions/ outcomes in respect of a set of information and finding an average of all such outcomes, which improves the accuracy of the prediction. Even for larger datasets, Random Forest predicts the output with high accuracy and very efficiently. It also helps to exterminate the limitations of a decision tree algorithm by reducing the modelling error in statistics and increasing precision.

$$\mathrm{RF}fi_i = \frac{\sum_{j \in alltrees} normfi_{ij}}{T}$$

Where, $RFfi$ denotes the feature i's importance, n$ormfi$ denotes the normalized feature importance of feature i in ij tree j T denotes the total number of trees.

Ridge Regression

It is a model tuning method used to analyse any data plagued by multicollinearity. It is the most common type of strategy used to deal with overpopulation (modelling error in mathematics). Ridge retrieval is a standard tuning method that can be used to analyse data with multicollinearity. The standard L2 method is used here. When there is a problem with multicollinearity, the least-squares are not biased, and variability is important, leading to estimated values far from real values.

Ridge retrieval cost function is given by

Min (|| Y - X (θ) || ^ 2 + λ || θ || ^ 2)

The time for punishment is lambda. The alpha function of the ridge function defines the value given here. We can control the time of punishment by changing the alpha values. The higher the alpha value, the greater the penalty, and as a

result the coefficient size is reduced.

It reduces the size of the parameters. As a result, it's applied to avoid multicollinearity.

By shrinking the coefficients, it minimises the model's complexity.

RESULTS & DISCUSSIONS

Machine learning models of installed equipment are very complex by nature and their meaning and accuracy are explained by a complete error meaning and a validation verification. Trading Bias Variance is an important aspect of the entire research paper. The remains are traced to two locations. Each individual variation has its corresponding value and contribution to statistical metrics and is best illustrated with a combination of line chart and bar chart that reflects Root Mean Square Error and Mean Absolute Error. The chart below shows the machine learning model capabilities and their statistical metrics.

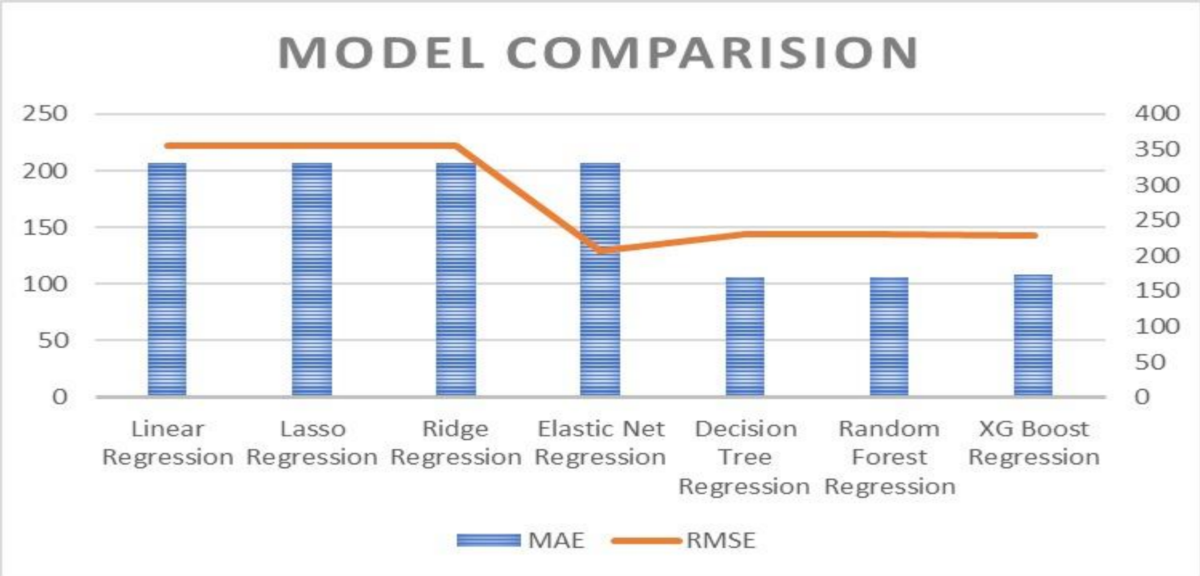

Figure 6: Model Comparison using MAE and RMSE Value

CONCLUSIONS

As the world's population grows, so does the need for food, and in recent years the number of people suffering from hunger, even after the greatest famine, is increasing daily. Governments and organizations working in the food industry plan and prepare solutions to prevent problems that may arise along the way from food safety for future generations. In this research paper, comparison with traditional mathematical techniques is demonstrated for better sales forecast, as shown by higher statistical metrics. There are older machine learning models with greater handling architecture and flexibility for data variables resulting in high power consumption and data volumes. Research has shown that complex problems, such as impact effect, bias variance trade off, multicollinearity are better handled by Machine Learning solutions. One concern about adopting Machine Learning for demand prediction can be a variety of available algorithms, thus making a unique choice hard. New research on actual implementation cases as well the methods used can help in the corresponding situation. Thus, this research contributes to the identification of benefits and features of machine learning, applied to

improve the accuracy of the demand forecast on FCMG industry, which is an important aspect of competition

REFERENCES

1. *Arvan, M., Fahimnia, B., Reisi, M., Siemsen, E. (2018). Integrating Human Judgement into Quantitative Forecasting Methods: A Review. Omega*
2. *Ferreira K.J., Lee B.H.A., Simchi-Levi D. (2016) Analytics for an online retailer: Demand forecasting and price optimization. Institute for Operations Research and the Management Sciences, v. 18, n. 1, 69-88.*
3. *Kandananond, K. (2012). A comparison of various forecasting methods for autocorrelated time series, International Journal of Engineering Business Management.*
4. *Liu, N; Ren, Sy; Choi, Tm; Hui, Cl; Ng, Sf. (2013). Sales Forecasting for Fashion Retailing Service Industry: A Review. Mathematical Problems in Engineering*
5. *MichalskI, R. S., Carbonell, J. G. and Mitchell, T. M. (1983). Machine Learning. Tioga publishing company, Palo Alto*
6. *Soguero-Ruiz C., Gimeno-Blanes F.-J., Mora-Jiménez I., Martínez-Ruiz M.P., Rojo-Álvarez J.-L. (2012). On the differential benchmarking of promotional efficiency with machine learning modelling (II): Practical applications. Expert Systems with Applications.*
7. *Tehrani, Af., Ahrens, D. (2016). Enhanced predictive models for purchasing in the fashion field by using kernel machine regression equipped with ordinal logistic regression. Journal of Retailing and Consumer Services.*
8. *Teucke, M., Ait-Alla, A., El-Berishy, N., Beheshti-Kashi, S., Lutjen, M. (2016). Forecasting of Seasonal Apparel Products. Dynamics in Logistics, LDIC, Springer International Publishing AG*